World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:3, No:2, 2009

A Neural Network Approach in Predicting the Blood Glucose Level for Diabetic Patients

Zarita Zainuddin, Ong Pauline and Cemal Ardil

Abstract**—**Diabetes Mellitus is a chronic metabolic disorder, where the improper management of the blood glucose level in the diabetic patients will lead to the risk of heart attack, kidney disease and renal failure. This paper attempts to enhance the diagnostic accuracy of the advancing blood glucose levels of the diabetic patients, by combining principal component analysis and wavelet neural network. The proposed system makes separate blood glucose prediction in the morning, afternoon, evening and night intervals, using dataset from one patient covering a period of 77 days. Comparisons of the diagnostic accuracy with other neural network models, which use the same dataset are made. The comparison results showed overall improved accuracy, which indicates the effectiveness of this proposed system.

Keywords**—**Diabetes Mellitus, principal component analysis, time-series, wavelet neural network.

I. Introduction

DIABETES Mellitus (DM) is a chronic and progressive metabolic disorder, where according to the World Health Organization there are approximately 171 million people in this world suffering from diabetes. The number of diabetic patients is expected to increase by more than 100% by the year 2030 [1].

Common manifestations of diabetes are characterized by insufficient insulin production by pancreas, ineffective use of the insulin produced by the pancreas or hyperglycemia. Causes like obesity, hypertension, elevated cholesterol level, high fat diet and sedentary lifestyle are the common factors that contribute to the prevalence of diabetes. Development of renal failure, blindness, kidney disease and coronary artery disease are types of the severe damage which are resulted by improper management and late diagnosis of diabetes. Even though there is no established cure for diabetes, indeed, the blood glucose level of diabetic patients can be controlled by well-established treatments, proper nutrition and regular exercise [1]-[3].

Treatment program for a diabetic patient usually involves several times of insulin dose injection per day. Besides, based on the advice of a physician, self-monitoring of blood glucose level with a blood glucose measuring advice is also done by the patient himself. Incorporating the information like previous insulin injection dose, previous blood glucose measurement and meal, the modification to the dietary, exact time and dose of insulin injection can be determined by the patient, with the aid of a measuring device.

Currently, there are a number of different computer assisted approaches that have been derived for the self-monitoring of blood glucose level, which consists of compartment model [4], algorithmic model [5]-[6] or mathematical model [7]-[8]. However, the full picture of the blood glucose metabolism is much more complicated. There are many detectable factors like food intake and physical exercise (Table I) and undetectable factors that characterize the blood glucose control. The complexity of the metabolism and the uncertainty associated with the measurement of the factors make the mentioned approaches to be less accurate and more uncertain. Hence, we seek for an alternative prediction tool which is more reliable in predicting the blood glucose level.

In this paper, we propose an expert system which is based on principal component analysis (PCA) and wavelet neural network (WNN) with different embedded wavelet families in the hidden layer (Mexican Hat, Gaussian wavelet and Morlet) in the prediction of blood glucose concentration. Since the interactions between the factors for glucose metabolism are complex, multidimensional, highly nonlinear, chaotic, stochastically and time variant time-series, the neural network model seems to be a more suitable predictor, where it can model the input-output behavior of the glucose metabolism, without knowing the involved explicit internal processes. In fact, neural network models have been applied widely in predicting time-series data [9]-[11].

Different types of neural network models have been used in modeling the blood glucose metabolism, such as the back propagation neural network [12]-[17], recurrent neural network [16], [18], neural fuzzy system [19]-[21], Bayesian neural network [22], multilayer perceptrons (MLP) [23]-[26], polynomial network [24], radial basis function neural network (RBFNN) [23], [27], Elman neural network [18] and time-series convolution neural network [26] have been

TABLE I
DETECTABLE FACTORS THAT INFLUENCE THE BLOOG GLUCOSE LEVEL

Increased Blood Glucose Level	Decreased Blood Glucose Level
Food Intake	Exercise
Infection	Stress
Obesity	Insulin
Inactivity	Low Food Intake

This work was supported in part by the Ministry of Science Technology and Innovation (MOSTI) e-Science Grant under Grant 01-01-05-SF0243.

Zarita Zainuddin is with the School of Mathematical Sciences, University Science of Malaysia, 11800 Penang, Malaysia (phone: 60-046533940; fax: 60-046570910; e-mail: zarita@cs.usm.my).

Ong Pauline is with is with the School of Mathematical Sciences, University Science of Malaysia, 11800 Penang, Malaysia (e-mail: ongpauline930@yahoo.com).

Cemal Ardil is with National Academy of Aviation, AZ1045, Baku, Azerbaijan.

applied widely in addressing the prediction of blood glucose concentration. However, the application of WNN in modeling the blood glucose metabolism has never been studied by other researchers. Indeed, we would see later that the nature of WNN makes it a suitable tool for the forecasting of the time-series for the blood glucose metabolism.

This paper is organized as follows. In section II, the used methodology of the proposed expert system and materials are discussed, followed by the simulation result of the proposed system in blood glucose concentration prediction in section III. Performance comparison between the proposed system and the other neural network models which used the same dataset as this paper is also made in section III. Finally, conclusions and future work are given in section IV.

II. Material and Methodology

A. Data Acquisition

The data is provided by Kok [25], which covers a continuous period of 77 days from one patient. At each day, the patient will need to fill in all the required information in a form, shown in Table II.

As shown in Table II, a day is split into eight measurement points: night (NT), before breakfast (BB), after breakfast (AB), before lunch (BL), after lunch (AL), before dinner (BD), after dinner (AD) and before sleeping (BS). The measurement points are further categorized into four different intervals: morning, afternoon, evening and night, where each interval consists of three measurement points, which are, at the start of interval, during the interval and at the end of the interval (for example: before breakfast, after breakfast and before lunch).

The patient needs to fill in the information for time of glucose measurement, blood glucose concentration, dose of short acting insulin injection, dose of long acting insulin injection, food intake, stress and exercise at each time of data recording. The scale of one to five is used for the recording of stress and exercise, where the value of one indicates no exercise at all and very relaxing, whereas the value of five means heavy exercise and heavy stress.

TABLE II
SAMPLE OF THE DAILY INFORMATION FOR A PATIENT TO FILL IN

9 April 2004

Friday	NT	BB	AB	BL	AL	BD	AD	BS
Time		8:27		12:25		17:29		23:42
Glucose Level		7.8		6.6		7.4		8.3
Short Act. Insulin		7		7		9		0
Long Act. Insulin		22	0	0	16	0	30	3
Food		81		93		92		8
Exercise		3		2		3		2
Stress		2		2		2		2

NT-night, BB-before breakfast, AB-after breakfast, BL-before lunch, AL-after lunch, BD-before dinner, AD-after dinner, BS-before sleep, Act-Acting

Since the units for these variables are not the same, hence the data are scaled into mean of 0 and standard deviation of 1 before the prediction of glucose concentration is done by the neural network.

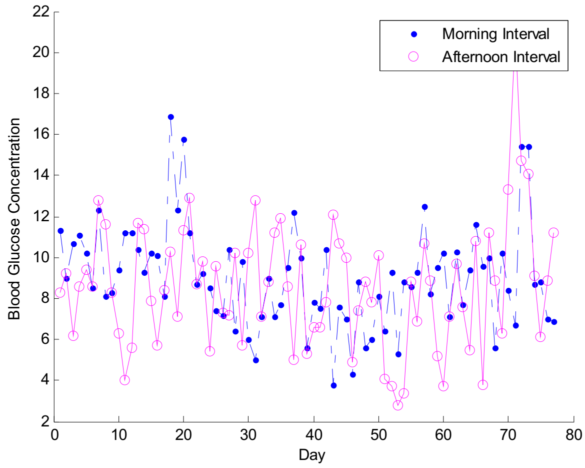

Fig. 1. Blood glucose concentration of a patient for a period of 77 days for the morning and afternoon interval. The unit for the glucose concentration is mmol/l. The connecting lines between the measurement points do not contribute any value

An example of a time-series for the blood glucose concentration for the morning and afternoon intervals is shown in Fig. 1. It can be observed that the time-series for the blood glucose concentration is highly non-stationary, time variant and chaotic. Hence, a neural network model appears to be a suitable tool to capture the behavior between the input-output dimensions.

The correlation coefficients, R, between the blood glucose concentration of morning, afternoon, evening and night intervals, calculated by using formula (1) are shown in Table III,

$$R = \frac{\mathrm{cov}(x, y)}{\mathrm{std}(x)\mathrm{std}(y)}, -1 \le R \le 1 \qquad (1)$$

where $\mathrm{cov}(x, y)$ is the covariance between variable x and y, $\mathrm{std}(x)$ and $\mathrm{std}(y)$ are the standard deviation of variable x and variable y respectively. From the statistical analysis of correlation coefficient, $R = 1$ indicates an increasing linear relationship, whereas $R = -1$ indicates a decreasing linear relationship between variable x and y. The values between the range $-1 \le R \le 1$ measure the degree of linear dependence between the variables. The variables are said to be highly correlated to each other if $R \ge \pm 0.5$.

From Table III, obviously, the blood glucose concentrations for the morning, afternoon, evening and night intervals are not highly correlated to each other. For example, a correlation coefficient of 0.1168 is obtained between the morning and afternoon intervals. This observation is probably due to the fact that the factors which control the glucose metabolism are dominating at different periods. The factors are dominating for a few hours, but are not continuous for the period that is longer than it. For example, the short acting insulin is injected before each meal in order to accommodate for the increased blood glucose concentration after the intake of food, but the effect of this short acting insulin can only stand for 3 to 4 hours. Thus, it is uncorrelated over a day period.

TABLE III
CORRELATION COEFFICIENTS BETWEEN THE BLOOD GLUCOSE CONCENTRATION FOR THE MORNING, AFTERNOON, EVENING AND NIGHT INTERVALS

	Morning	Afternoon	Evening	Night
Morning	1.0000	0.1168	0.1810	-0.0459
Afternoon	0.1168	1.0000	-0.0744	-0.1002
Evening	0.1810	-0.0744	1.0000	-0.1440
Night	-0.0459	-0.1002	-0.1440	1.0000

Therefore, based on the fact that the values for these four intervals are uncorrelated, separating the blood glucose prediction into four different neural network models is reasonable, where one neural network model is built for each interval.

B. Feature Selection

It has been proven that inclusion of the past measurement of the input variables increases the prediction accuracy of the neural network model considerably [18]. However, glucose metabolism process changes with time. Thus, the modeling of glucose metabolism which involves numbers of internal/external factors as well as past history of the factors itself is cumbersome. This is where the PCA plays its role, where it is capable of extracting the characteristic features of this complicated process.

Assume that along the time-series data, there are recurrent characteristic features with a scale of *L*. Moving a window of length *L* along the time-series with a step *s* at a time is comparable to finding the vectors of the corresponding dimension. Hence, those recurrent characteristic features will keep repeating in the window vector. This window vectors are the projected onto basis vectors through PCA. By applying PCA, the original window vectors which reflect the temporal variability of the blood glucose concentrations (fragment of time-series data) are transformed into a lower dimensionality of independent orthogonal principal components (PC). The PC will only correspond to the basic component of the characteristic features, where this implies that there is only one significant PC for those recurrent features, even though it repeats for more than once in the time-series. Hence, a highly irregular time-series data will be transformed into a new regular time-series of PC scores, which enables an easier prediction.

A brief explanation of PCA technique is given below [28].

Let $X_{m\times n}$ be the data matrix with *n* column vectors and *m* elements in each vector. Firstly, a covariance matrix, C, of the data matrix is calculated by using eqn. (2)

$$C_x = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})(x_i-\bar{x})^T \qquad (2)$$

where $\bar{x}$ is the mean of *X*.

Subsequently, the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ and the corresponding eigenvectors of the covariance matrix $E_1, E_2, \ldots, E_N$ are calculated. Then, the eigenvalues are sorted in decreasing order. Hence, one can create the *k*th PC, y_k by using the eqn. (3)

$$y_k = E_k^T X, k = 1,2,\ldots,N \qquad (3)$$

In fact, the PC is the linear combination of original data. The first PC represents the direction of maximum variation in the data, where the second PC, which is orthogonal to the first PC, represents the next largest variation in the data and so on. Since most of the variation in the data is concentrated in the first few PCs, hence choosing the first few PCs are accountable for most of the variability.

The number of PC to choose can be determined by using eqn. (4)

$$p = \frac{\lambda_1+\lambda_2+\lambda_3+\ldots+\lambda_w}{\lambda_1+\lambda_2+\lambda_3+\ldots+\lambda_N}\times 100\% \qquad (4)$$

where $w \le N$. A total of *w* PCs are chosen when the value of *p* exceeds a certain percentage.

In order to make the performance comparison, the same 19 input variables as Kok [25] and Baghdadi and Nasrabadi [27] are used, which are shown in Table IV. By applying PCA, the dimensionality of 19 variables is reduced to 4 PCs for the morning, afternoon, evening and night intervals respectively, since the first 4 PC already dominate for more than 90% variability. Subsequently, this first 4 PCs will be the input for the WNN model, which will be discussed in the next section, and the output of the WNN will be the predicted blood glucose concentration at the end of each interval.

C. Blood Glucose Level Prediction Based on Wavelet Neural Network

There are a number of different neural network models which differ in the network architecture, learning algorithm, number of hidden layers, and also the type of activation function used in the hidden layers. WNN is one of the neural network models which is inspired by the similarities between the wavelet decomposition and the single hidden layer neural network. Since the first implementation by Zhang and Benveniste [29]-[30], WNN have been applied widely in

TABLE IV
THE INPUT VARIABLES FOR THE PRECTION OF BLOOD GLUCOSE LEVEL

1.	Glucose Level	During Interval
2.	Short Acting Insulin	During Interval
3.	Food Intake	During Interval
4.	Exercise	During Interval
5.	Glucose Level	At Start of Interval
6.	Long Acting Insulin	During Past 24 Hours
7.	Stress	During Interval
8.	Glucose Level	At Start of Interval on Previous Day
9.	Short Acting Insulin	During Interval on Previous Day
10.	Food Intake	During Interval on Previous Day
11.	Exercise	During Interval on Previous Day
12.	Resulting Glucose	At End of Interval on Previous Day
13.	Glucose Level	At Start of Previous Interval
14.	Short Acting Insulin	During Previous Interval
15.	Food Intake	During Previous Interval
16.	Exercise	During Previous Interval
17.	Exercise	Average Over Past 24 Hours
18.	Interval Length	During Interval
19.	Exercise Added Up Squared Values	During Past 24 Hours

various applications, such as system identification, classification, and pattern recognition problems [31]-[33].

WNN differ from the other neural network models in that it involves the integration of wavelet families as the activation function in the hidden nodes. A schematic diagram of a WNN, with d input nodes, m hidden nodes and L output nodes is shown in Fig. 2.

The input layer will receive the input variable $\vec{x} = (x_1, ..., x_d)$ and transmit the accepted input variables to the next layer. The second layer is a hidden layer with a mother wavelet ψ in each hidden node. In this paper, different wavelet families, namely Mexican Hat, Gaussian wavelet and Morlet are selected as the basis functions in hidden layer. The nodes in this layer are given by the product of the mother wavelet as

$$\psi_j(\zeta) = \psi(\| D_j(X - t_j) \|), j = 1,...,m \tag{5}$$

where D_j and t_j are the scaling and translation vector respectively.

The sigmoid function (logistic and hyperbolic tangent) is the commonly used basis function in an MLP. Compared to the basis functions in the hidden layers of the MLP, the mother wavelet used in the hidden nodes of a WNN is a localized activation function. Hence, the connection weight associated with the hidden nodes can be viewed as locally piecewise constant models, which leads to learning efficiency and structure transparency.

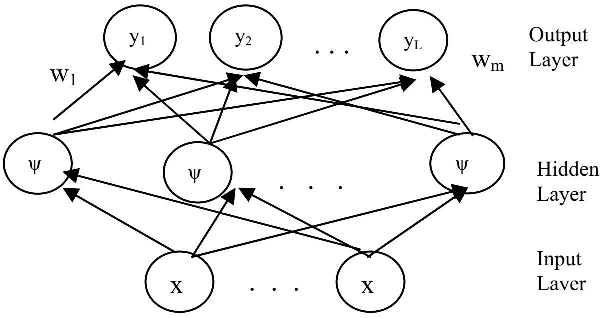

Fig. 2. A schematic diagram of a wavelet neural network with d input nodes, m hidden nodes and L output nodes

The third layer is the output layer. The output will be the linear combination of the weighted sum of the hidden layer, which is given by eqn. (6)

$$y_k(x) = \sum_{j=1}^{m} w_j \psi_j(\| D_j(X - t_j \|) + \Theta, k = 1,2,...,L \tag{6}$$

where w_j and Θ are the weight vector and bias term between hidden layer and output layer respectively. Obviously, all the neurons in any layer are fully connected to the preceding and also the succeeding layer, but no connections between the neurons within the same layer are allowed.

Various learning algorithms can be applied to the training of WNN; in this paper, the learning of WNN is by the method of solving the pseudo-inverse with fixed parameter initialization. Therefore, only the weight matrix W needs to be adjusted during the training of WNN, in order to map the underlying relationship between input and output space.

Before we begin to describe the learning algorithm of WNN, let us define the cost function as in eqn. (7)

$$E(f(n)) = \frac{1}{2}(y_d(n) - y(n))^2 \tag{7}$$

where y_d is the desired output value and $y(n)$ is the output value from WNN. Hence, the training of WNN is based on the minimization of the cost function.

There are two stages involved. Firstly, the scaling parameter is fixed. Next, the translation vector is randomly chosen from the input vectors. Let us represent eqn. (6) as $Y = \psi W$, where

$$\psi = \begin{pmatrix} \psi(x_1, D_1, t_1) & \psi(x_1, D_2, t_2) & \cdots & \psi(x_1, D_m, t_m) \\ \psi(x_2, D_1, t_1) & \psi(x_2, D_2, t_2) & \cdots & \psi(x_2, D_m, t_m) \\ \vdots & \vdots & \vdots & \vdots \\ \psi(x_d, D_1, t_1) & \psi(x_d, D_2, t_2) & \cdots & \psi(x_d, D_m, t_m) \end{pmatrix} \tag{8}$$

is the output of wavelet families ψ, and $\psi(x_i, D_i, t_i) = \psi(\| D_i(x_i - t_i) \|)$.

Therefore, in order to solve the weight matrix W, $W = \psi^+ Y$ is computed. ψ^+ is the pseudo-inverse defined as $\psi^+ = (\psi^T \psi)^{-1} \psi^T$.

A summary of the learning algorithm of WNN is given as below:

i. Initialize the values for D_i and t_i.
ii. Feed in the input vector X into the WNN.
iii. Calculate the product of the hidden layer by using eqn. (5).
iv. Solve for the weight matrix W by using the pseudo-inverse method.
v. Obtain the output value of WNN, $y(n)$ from step (iv).
vi. Compare $y(n)$ with the desired output value, y_d.
vii. Calculate the cost function as in eqn. (7).
viii. Repeat steps (ii) to (vii) until it meets the stopping criterion.

D. *Performance Index*

The training of the WNN model involves the modification of the weight vector gradually, in order to minimize the difference between the predicted value and the desired response. The difference is referred as cost function, which can be measured by various criteria.

In this paper, the performance index in terms of root mean squared error (RMSE) is used, which is formulated as

$$RMSE = \sqrt{\frac{1}{p}\sum_{j=1}^{p}[f(x_j) - y_j]^2} \tag{9}$$

where p is number of testing samples, $f(x_j)$ is the desired response and y_j is the predicted output of WNN.

E. *Multifold Cross Validation*

Excessive training will force the WNN to memorize the input vectors and insufficient training will cause the WNN to be unable to learn from the input vectors presented to it, where it will lead to poor generalization when new inputs are presented to the WNN.

Therefore, in order to avoid these problems, multifold cross validation is used. The samples are divided into k groups, where $k > 1$. Firstly, one group from the samples is left out, where the training of the neural network involves the remaining of the samples. Next, the validation error is measured by testing it on the group left out. The process is repeated for k times, using each different group respectively as the testing set. Subsequently, the average of the validation error is calculated. In this study, we use a 10-fold cross validation.

III. Experimental Results

As discussed in Section II, four different WNN models will be constructed separately for the morning, afternoon, evening and night intervals, in order to predict the blood glucose concentration at the end of each interval.

There are two stages involved in the prediction of the blood glucose concentration by using the proposed system. Firstly, the input dimension which consists of 19 input variables is reduced to 4 PCs by using the PCA technique. Subsequently, the PCs will be the input to the input layer of the WNN. After proper training, the output from the WNN will be the predicted blood glucose concentration at the end of the interval.

The experimental results (in terms of RMSE) of this proposed system by using different wavelet families in the hidden layer of WNN are shown in Table V.

A. Performance Assessment and Discussion

From Table V, it can be seen that all the proposed system with different wavelet families in the hidden nodes performed well in predicting the blood glucose concentration at end of each interval. The highest RMSE is at 0.0504 for the morning interval and the lowest RMSE is at 0.0170 for the night interval. These RMSEs are rather small, indicating a promising performance of the proposed system.

A plot of the time-series for the predicted and measured blood glucose concentrations at the end of morning interval by using the proposed expert system with Gaussian wavelet in hidden layer of WNN is shown in Fig. 3, where Fig.4 is the enlarged segment of Fig. 3, for the first 30 days. It can be seen that, a relatively good prediction can be obtained by using the proposed expert system.

From the experimental results, the proposed expert system with Gaussian wavelet in the hidden nodes of WNN produces the lowest RMSE for each interval, when it was compared with the WNN with Mexican Hat or Morlet in the hidden layer.

As shown in Fig. 1, the time-series for the blood glucose concentration is highly irregular. It varies considerably and it has a number of big leaps throughout the time-series. From the wavelet analysis, high intensity signal is produced when the data correlates strongly to the shape of the wavelet. Thus, using the wavelet families which have the shape that is identical to the shape of the features being investigated is crucial. In other words, the wavelet families that have a great change in its main feature will manage to detect the great changes in the time-series of blood glucose concentration.

TABLE V
PREDICTION OF BLOOD GLUCOSE LEVEL USING THE PROPOSED SYSTEM WITH DIFFERENT WAVELET FAMILIES IN THE HIDDEN LAYER OF WNN

Wavelet Families	RMSE for Different Intervals			
	Morning	Afternoon	Evening	Night
Mexican Hat	0.0460	0.0365	0.0370	0.0175
Gaussian Wavelet	0.0450	0.0348	0.0330	0.0170
Morlet	0.0504	0.0350	0.0369	0.0176

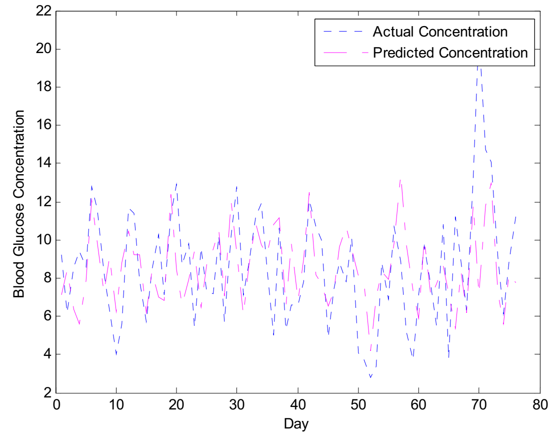

Fig. 3. A plot of time-series for the predicted and actual measured blood glucose concentration at the end of the morning interval by using the proposed system with Gaussian wavelet in the hidden layer of the WNN

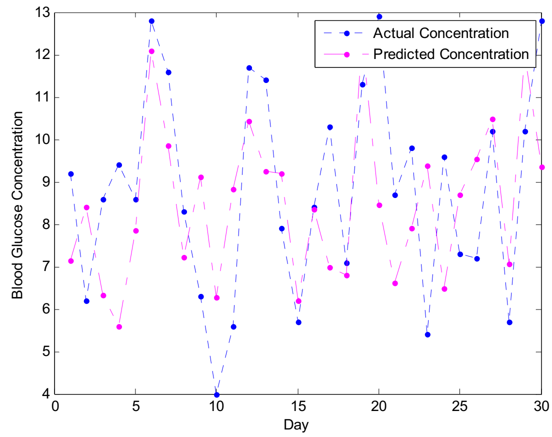

Fig. 4. An enlarged segment of the time-series in Fig. 3, for the first 30 days

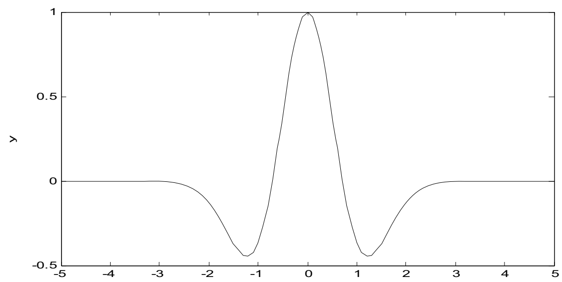

Fig. 5a. A schematic diagram for Mexican Hat

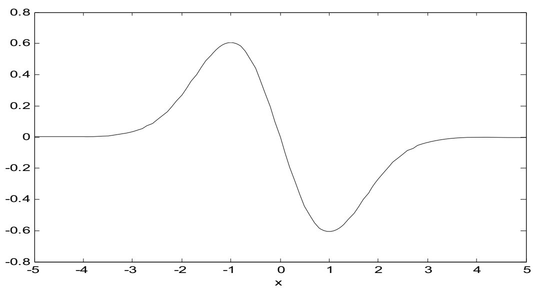

Fig. 5b. A schematic diagram for Gaussian wavelet

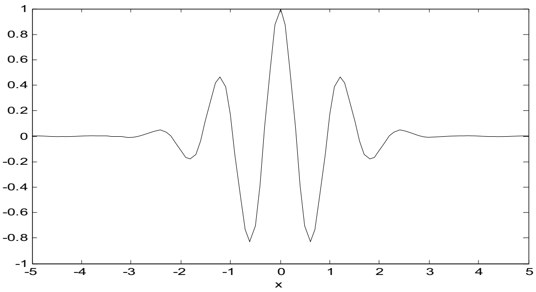

Fig. 5c. A schematic diagram for Morlet

The schematic diagram of the wavelet families which we used in this study, namely Mexican Hat, Gaussian wavelet and Morlet, is given in Fig. 5. These wavelet families are symmetric (Mexican Hat and Morlet) or anti-symmetric (Gaussian wavelet) and they are highly regular. Wavelet families of Mexican Hat and Morlet work well in approximating the quasi-sinusoidal function [34].

However, from Fig. 1, the time-series is of a saw-tooth shape, instead of sinusoidal. Therefore, the wavelet families that have a large change in its shape will be more suitable in predicting the saw-tooth time-series, where Gaussian wavelet satisfies this criterion. The main feature of this anti-symmetric Gaussian wavelet might explain the good prediction accuracy for the saw-tooth time series that are used in this study.

The present glucose metabolism is affected by the past glucose metabolism. That is the reason why the input variables from the previous interval are also considered in modeling the glucose metabolism in a particular interval (Table IV). However, the blood glucose concentration at a given instant, even though depending on the past history of the input variables, should only be limited within a certain short time interval. When the period of time between two consecutive measurements of the blood glucose level are getting longer, the influence of the history of the variables from the previous interval are not significantly important anymore. Somehow, this history information might become useless and add noise to the data. The length of the time interval between two consecutive measurements for the night interval and the morning interval of the following day spans a period of around 8 hours, which is a long period of time. This might explain the lowest accuracy of the blood glucose prediction for the morning interval.

B. Performance Comparison and Discussion

Performance comparison between the proposed expert system with the other neural network models, namely, MLP [25] and RBFNN [27] are made, using the same dataset in earlier studies. The same 19 input variables (Table IV) are considered, but different input selection approaches are applied.

Kok [25] used MLP in predicting the blood glucose levels, where a heuristic approach is used in selecting the input variables for the training of the neural network model. By trial and error, Kok came out with 12 different combinations of the input variables, and the MLP were trained by these different input combinations for each morning, afternoon, evening and night interval. Performance criteria are calculated from the testing samples, where the particular input selection method that performed the best for each interval was selected, followed by the blood glucose levels prediction by using an MLP.

As a result from this heuristic approach, Kok used input variables (2, 3, 4, 5, 6, 7, 18, 19) for the morning and night interval, (2, 3, 4, 5, 7) for the afternoon interval and (2, 5, 3, 4, 13, 14, 15, 16) for the evening interval (Please refer Table IV for the information on input variables).

Followed from Kok, Baghdadi and Nasrabadi [27] used an RBFNN in modeling the blood glucose metabolism, where a pruning method was applied in the input selection. Firstly, the RBFNN was trained by all the 19 input variables. After the training of RBFNN, the criterion function as well as the weight of each input variable was calculated. The input variable with a low magnitude of weight was eliminated, followed with the training of RBFNN again. The criteria functions were measured again. Elimination of the input variable is worthy, if the present criterion functions were better than the previous one. Otherwise, the eliminated input variable will not be omitted. The process was repeated until the optimal input variables for each interval were obtained.

In the end, after the elimination of the variables, input selection of (1, 2, 5, 6, 7, 10, 15, 17, 18) for morning interval, and (1, 2, 5, 10, 15, 18) for afternoon, evening and night interval are obtained (Please refer Table IV for the information of input variables).

The performance comparison between these different input selection methods and predictor based on different neural network models is given in Table VI. The performance of the proposed expert system with Gaussian wavelet in the hidden layer of WNN is used for the comparison.

From Table VI, a combination of input selection by heuristic approach and predictor based on MLP used by Kok performed unsatisfactorily. This disappointing performance is probably due to the poor input selection and also the behavior of the MLP. Selecting the combination of input variables randomly might overlook the implicit relationship between the variables since there are still many undetectable relationships between the factors that affect the blood glucose metabolism. Hence, selecting the input variables by trial and error seems to be not a wise approach.

Besides that, the sigmoid function used as the activation functions in the MLP is a globalized activation function. Compared with the localized wavelet functions in the hidden layer of WNN which leads to learning efficiency, the learning of the MLP is time consuming. In addition, the back propagation learning algorithm used by MLP tends to get trapped in local minima. When the MLP is unable to converge to a global minimum, the prediction accuracy of the MLP will deteriorate. Thus, this might explain the poor performance of Kok's result.

From Table VI, the performance of the proposed system outperforms the RBFNN used by Baghdadi and Nasrabadi [27] overall.

Regardless of the input selection approach, when we only emphasized on the predictive capability of RBFNN and WNN, the latter seems to be a more appropriate predictor for the time-series data. This is mainly due to the activation function used in the hidden layer of RBFNN and WNN. In [4], the Gaussian function is used in the hidden nodes of RBFNN. However, a periodic function and an exponential function are approximated better by using a WNN with an oscillating wavelet basis function and by a Gaussian activation function respectively. When we look at the nature of the time-series of the blood glucose concentration, approximating its saw-tooth behavior is more suited by using WNN. This might justify the superior performance of the proposed system for the morning, afternoon and evening intervals, when its predicting accuracies are compared against with the performance of RBFNN.

For the night interval, a RMSE of 0.0118 is obtained for the RBFNN, whereas for the proposed system, a RMSE of 0.0170 is achieved. In this interval, the performance of RBFNN is more superior. This might be caused by the input selection method. When we look at the selected input variables for the night interval used by the RBFNN approach, the number of factors in the previous interval, previous day or the past 24 hours are eliminated. Compared with the PCA technique that we used for the input selection, the influence for all the factors that play an important role in blood glucose metabolism, as well as the history for all the factors are taken into account in the PCA. However, even though the blood glucose concentration at a given instant depends on its past history, when the length of the time interval is too long, the past history of the input variables on the previous day or past 24 hours might not influence the present blood glucose concentration anymore. We should clarify that when we mention about the term past 24 hours, it indicates the time interval start from 00:00am to 24:00pm on the previous day. Hence, the length of the time interval between the measurement points for the input variables during the past 24 hours is longer for the night interval, compared with the morning, afternoon and evening intervals. Hence, when we take it into the consideration by using the PCA, it might deteriorate the predictive capability of the WNN.

IV. Conclusion

Elevation of blood glucose level abruptly will make the diabetic patients go into a coma. Hence, prediction of the blood glucose concentration is important, in order for the patients to adjust the dose for insulin injection and for preventing the severe complications that result from improper management of the blood glucose levels.

TABLE VI
PERFORMANCE COMPARISON BETWEEN THE PROPOSED SYSTEM AND THE OTHER NEURAL NETWORK MODELS

	RMSE for Different Intervals			
	Morning	Afternoon	Evening	Night
Kok [18]	1.8	1.8	2.1	2.2
Baghdadi and Nasrabadi [4]	0.0826	0.0513	0.0373	0.0118
The Proposed System	0.0450	0.0348	0.0330	0.0170

In this paper, we have proposed an expert system, where a feature selection based on the PCA and a predictor based on WNN are used. Our experimental results showed that the proposed expert system is a powerful model for the blood glucose prediction and it outperformed both the MLP and RBFNN approach. The better performance of the proposed system might be attributed to the integration of wavelet families in the hidden layer of WNN, which can be explained by the fact that wavelet functions can capture the behavior of the chaotic time-series better, compared with the sigmoid function in an MLP and the Gaussian function in a RBFNN.

The promising experimental results are also probably contributed by the input selection by using PCA, where the influences of the input variables at present as well as its past history are taken into consideration of the blood glucose metabolism. However, there is an issue related to the time constant. Some of the past history of the input variables that we have considered in the PCA, span for more than 36 hours from the present measurement point. The question that arises here is, how far behind from the current measurement point should we take into account.

In future work, besides the issue of the time constant for the past history of the input variables, the other significant factors that influence the glucose metabolism, like weight, age and sex, as well as the physiological parameters such as heart rate and skin impedance can be integrated in the modeling of the neural network. Development of a system that involves the prediction of the blood glucose level, as well as recommendation for the appropriate therapy is favorable.

Besides that, in this study, we have only emphasized on the modification of the wavelet families in the hidden layer of WNN. The parameter initialization and the learning algorithm of WNN are in primitive form. In future, advanced parameter initialization method and learning algorithm, by using clustering, genetic algorithm, particle swarm optimization can be applied in improving the performance of WNN. Integration of WNN with fuzzy logic is another issue that can be pursued.

Acknowledgment

The authors wish to express their gratitude to Peter Kok from Delft University of Technology for providing the diabetes data and his valuable discussion, and also Golnaz Baghdadi and Ali Motie Nasrabadi from Shahed University for their discussions.

References

[1] World Health Organization. Available: http://www.who.int

[2] American Diabetes Asscociation. Available: http://www.diabetes.org

[3] Gan, D. editor. Diabetes atlas, 2nd ed. Brussels: International Diabetes Federation, 2003. Available at http://www.eatlas.idf.org/webdata/docs/Atlas%202003-Summary.pdf

[4] C.V. Doran, N.H. Hudson,K.T. Moorhead, J.G. Chase, G.M. Shaw, and C.E. Hann, "Derivative weighted active insulin control modeling and clinical trials for ICU patients," *Medical Engineering and Physics*, vol. 26, pp. 855-866, 2004.

[5] P. Dua, F.J. Doyle III, E.N. Pistikopoulos, "Model-based blood glucose control for type I diabetes via parametric programming," *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 1478-1491, 2006.

[6] C. Owens, H. Zisser, L. Jovanovic, B. Srinivasan, D. Bonvin, and F.J. Doyle III, "Run-to-run control of blood glucose concentrations for people with type I diabetes mellitus," *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 996-1005, 2006.

[7] P.G. Fabietti, V. Canonico, M.O. Federici, M. Massimo, and E. Sarti, "Control oriented model of insulin and glucose dynamics in type I

diabetics," *Medical and Biological Engineering and Computing*, vol. 44, pp. 69-78, 2006.

[8] M.S. Ibbini, M.A. Masadeh, and M.M.B. Arner, "A semiclosed-loop optimal control system for blood glucose level in diabetics," *Journal of Medical Engineering and Technology*, vol. 28, pp. 189-196, 2004.

[9] B.R. Chang, and H.F. Tsai, "Forecast approach using neural network adaption to support vector regression grey model and generalized auto-regressive conditional heteroscedasticity,: *Expert Systems with Application*, vol. 34, pp. 925-934, 2008.

[10] M. Engin, S. Demirag, E.Z. Engin, G. Celebi, F. Ersan, E. Asena, and Colakoğlu, "The classification of human tremor signals using artificial neural network," *Expert Systems with Applications*, vol. 33, pp. 754-761, 2007.

[11] M. Ture, and I. Kurt, "Comparison of four different time series methods to forecast hepatitis A virus infection," *Expert Systems with Applications*, vol. 31, pp. 41-46, 2006.

[12] A.K. El-Jabali, "Neural network modeling and control of type I diabetes mellitus," *Bioprocess Biosystem Engineering*, vol. 27, pp. 75-79, 2005.

[13] T.N. Hung, G. Nejhded, and W.J. Timothy, "Neural-network detection of Hypoglycemic episodes in children with type I diabetes using physiological parameters," *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York, 2006, pp. 6053-6056.

[14] C. Li, and R. Hu, "PID control based on BP neural network for the regulation of blood glucose level in diabetes," *Proceedings of the 7th International Conference on Bioinformatics and Bioengineering*, Boston, pp. 1168-1172, 2007.

[15] J.J. Liszka-Hackzell, "Prediction of blood glucose levels in diabetic patients using a hybrid AI technique," *Computers and Biomedical Research*, vol. 32, pp. 132-144, 1999.

[16] S.G. Mougiakakou, A. Prountzou, D. Iliopoulou, K.S. Nikita, A.Vazeou, and C.S. Bartsocas, "Neural network based glucose-insulin metabolism models for children with type I diabetes," *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York, 2006, pp. 3545-3548.

[17] K. Zarkogianni, S.G. Mougiakakou, A. Prountzou, A. Vazeou, C.S. Bartsocas, and K.S. Nikita, "An insulin infusion advisory system for type I diabetes patients based on non-linear model predictive control methods," *Proceedings of the 29th IEEE EMBS Annual International Conference*, Lyon, 2007, pp. 5971-5974.

[18] V. Tresp, T. Briegel, and J. Moody, "Neural-network models for the blood glucose metabolism for a diabetic," *IEEE Transactions on Neural Networks*, vol. 10, pp. 1204-1213, 1999.

[19] D. Dazzi, F. Taddei, A. Gavarini, E. Uggeri, R. Negro, and A. Pezzarossa, "The control of blood glucose in the critical diabetic patient: A neuro-fuzzy approach," *Journal of Diabetes and Its Complications*, vol. 15, pp. 80-87, 2001.

[20] Phee, H.K., Tung, W.L. & Quek, C. (2007). A personalized approach to insulin regulation using brain-inspired neural semantic memory in diabetic glucose control. *IEEE Congress on Evolutionary Computation*, Singapore, pp. 2644-2651.

[21] K. Polat, and S. Günes, :An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease," *Digital Signal Processing*, vol. 17, pp. 702-710, 2007.

[22] T.N. Hung, G. Nejhded, T.N. Son, and W.J. Timothy, "Detection of hypoglycemic episodes in children with type I diabetes using an optimal Bayesian neural network algorithm," *Proceedings of the 29th IEEE EMBS Annual International Conference*, Lyon, 2007, pp. 3140-3143.

[23] R. Abu Zitar, "Towards neural network model for insulin/glucose in diabetics," *International Journal of Computing & Information Sciences*, vol. 1, pp. 25-32, 2003.

[24] R. Abu Zitar, and A. Al-Jabali, "Towards neural network model for insulin/glucose in diabetics-II.," *Informatica*, vol. 29, pp. 227-232, 2005.

[25] P. Kok, "*Prediction blood glucose levels of diabetics using artificial neural networks*," Research Assignment for Master of Science, Delft University of Technology, 2004.

[26] S.A. Quchani, and E. Tahami, "Comparison of MLP and Elman neural network for blood glucose level prediction in type I diabetics," *Proceedings of the 3rd International Federal of Medical and Biological Engineering*, Kuala Lumpur, 2007, pp. 54-58.

[27] G. Baghdadi, and A.M. Nasrabadi, "Controlling blood glucose levels in diabetics by neural network predictor," *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, Lyon, 2007, pp. 3216-3219.

[28] A.J. Richard, and W.W. Dean, *Applied multivariate statistical analysis*. New Jersey: Prentice-Hall, 2002, ch 4.

[29] Q. Zhang, "Using wavelet network in nonparametric estimation," *IEEE Transactions on Neural Networks*, vol. 8, pp. 227-236, 1997.

[30] Q. Zhang, and A. Beveniste, "Wavelet Networks," *IEEE Transactions on Neural Networks*, vol. 3, pp. 889-898, 1992.

[31] B. Biswal, P.K. Dash, B.K. Panigrahi, and J.B.V. Reddy, "Power signal classification using dynamic wavelet network," *Applied Soft Computing*, vol. 9, pp. 118-125, 2009.

[32] S. Srivastava, M. Singh, M. Hanmandlu, and A.N. Jha, "New fuzzy wavelet neural networks for system identification and control," *Applied Soft Computing*, vol. 6, pp. 1-17, 2005.

[33] H. Zhang, B. Zhang, W. Huang, and Q. Tian, "Gabor wavelet associated memory for face recognition," *IEEE Transactions on Neural Networks*, vol. 16, pp. 275-278, 2005.

[34] J.C. Hargreaves, "Timing of ice-age terminations determined by wavelet methods," *Paleoceanography*, vol. 18, pp. 1035-1048, 2003.